Reduction methods for qualitive and quantative data in regional analysis

Brouwer, F.; Nijkamp, P.

1982

document version
Publisher's PDF, also known as Version of record

Reduction Methods for Qualitative and Quantitative Data in Regional Analysis

Floor Brouwer
Peter Nijkamp

August 1982

Researchmemorandum 1982-21

Introduction

In many multivariate methods the underlying variables are supposed to be measured on a cardinal scale. In these cases, the variables and concepts are quantified on either a ratio-scale or an interval-scale, and numerical operations like addition and multiplication, are permitted. But we have to employ other methods when variables are either inaccurate or measured on a qualitative scale. There are different methods to manipulate qualitative data in a correct way that lead to cardinal or metric information. This set of techniques is developed with two aims:

- reduction of a data-set with non-metric observations. The original data input will be represented in smaller dimensions, depending on a "goodness-of-fit statistic" or a previously chosen classification criterium.
- transformation of the data on a cardinal scale.

Three reduction methods are developed in this paper, i.e. ordinal principal component analysis, qualitative cluster analysis and a multidimensional scaling method (or "geometric scaling method").

These methods will be illustrated with Dutch regional data of socio-economic, environmental and infrastructural profile elements.

2. Principal Component Analysis

One of the possibilities for measuring the agreement between attributes of areas is the use of principal component analysis (PCA). The aim of PCA is the representation of the original J variables by a reduced set of, say k, other (independent) variables, where these variables have a high degree of correlation with the J variables. This type of data reduction is often suggested as a solution to the multicollinearity problem. By means of PCA one may obtain insight into the linkage between both sets and the way of determining variables, based on an orthogonal data transformation which will be explained below. A theoretical description of PCA and a determination of the formulae are found in Nijkamp and Paelinck, 1976; Maddala, 1977; Theil, 1979. A practical foundation and interpretation of results is given in Hauer and Van der Knaap, 1973 and Nijkamp and Rietveld, 1981.

When, for example, all variables would be proportionally related to each other, we would have a set of maximal correlated observations i.e. with correlation values equal to one. In this case, a principal component analysis, leading to a set of uncorrelated factors, would be very easy. One single variable would suffice to describe the behaviour of all J variables.

A principal component analysis is only a redefinition of original observations by means of an orthogonal data transformation. From this follows that the number of factors obtained from a principal component analysis is, in principle, exactly equal to the number of original variables.

Suppose a data matrix N, of order I x J, with I observations for all J variables. Then we can consider linear functions of these variables, say :

$$\begin{aligned} p_1 &= a_1 n_1 + a_2 n_2 + \ldots + a_J n_J \\ p_2 &= b_1 n_1 + b_2 n_2 + \ldots + b_J n_J \text{ , etc.} \end{aligned} \qquad (1)$$

The elements n_j , j=1, ..., J correspond to the observations on variable j.

The vectors a_j are chosen such that $var(p_1)$ is maximized subject to the normalization condition $a_1^2 + \ldots + a_J^2 = 1$. Now, p_1 is called the first principal component; it has the highest variance under the given normalization restriction. This condition is necessary to avoid indefinitely large variances. The second principal component p_2, to be uncorrelated with p_1, has maximum variance subject to the condition $b_1^2 + \ldots + b_J^2 = 1$. For all J possible principal components $p_1, p_2, \ldots, p_J$ it can be shown that

$$var(p_1) + var(p_2) + \ldots + var(p_J) = var(n_1) + var(n_2) + \ldots + var(n_J) \qquad (2)$$

While the principal components are mutually orthogonal or uncorrelated, the elements $n_1, n_2, \ldots, n_J$ may be highly correlated.
By means of (1) it follows that the index p is a linear combination of the variables, rewritten as :

$$\underline{p} = N \underline{a} \text{ ,} \qquad (3)$$

where $\underline{p}$ and $\underline{a}$ are vectors of length I and J respectively, and N is a data matrix of order I x J. From (3) follows

$$(N' N)^{-1} N' \underline{p} = \underline{a} \qquad (4)$$

1) N' is the transpose of matrix N , and hence of order J x I .

Formula (4) will be used below. The method of PCA aims to represent N by means of $\underline{p}\ \underline{a}'$ with a high degree of accuracy. This means that the matrix N consisting of IJ parameters is approximated by the matrix $\underline{p}\ \underline{a}'$ when only one component is defined, and which is determined by I + J parameters. The matrix of deviations between the original matrix N and the approximated one, $\underline{p}\ \underline{a}'$, is given by:

$$D = N - \underline{p}\ \underline{a}' = N - N\ \underline{a}\ \underline{a}' = N(I - \underline{a}\ \underline{a}') \tag{5}$$

The original matrix N may correspond rather well with the approximated one, $\underline{p}\ \underline{a}'$, when the sum of squares of deviations defined by (5) is minimized, i.e.

$$\underset{\underline{a},\underline{p}}{\text{Min}} : \sum_{i=1}^{I} \sum_{j=1}^{J} d_{ij}^2 \tag{6}$$

The range of possible values of the vector $\underline{a}$ is restricted by means of a normalization, such that

$$\underline{a}'\ \underline{a} = 1 \tag{7}$$

The minimization procedure given by (6) is the same as the determination of the maximum value of the trace from the matrix D D'. The trace can be rewritten as

$$\begin{aligned} \text{tr}(D\ D') &= \text{tr}(\ N(I - \underline{a}\ \underline{a}')\ N'(I - \underline{a}\ \underline{a}')) \\ &= \text{tr}\ N\ N' - \text{tr}\ N\ \underline{a}\ \underline{a}'\ N' \quad {}^{1)} \\ &= \text{tr}\ N'\ N - \underline{a}'\ N'\ N\ \underline{a}. \quad {}^{2)} \end{aligned} \tag{8}$$

The trace of the matrix N' N is determined completely from the observations, and is independent from $\underline{a}$ and $\underline{p}$. From this follows that minimization of (8) will be the same as maximization of $\underline{a}'\ N'\ N\ \underline{a}$. This can be written by means of a Lagrange function, which has to be maximized:

$$L = \underline{a}'\ N'\ N\ \underline{a} - \lambda\ (\underline{a}'\underline{a} - 1) \tag{9}$$

1) $(I - \underline{a}\ \underline{a}')$ is idempotent, which means $(I - \underline{a}\ \underline{a}')(I - \underline{a}\ \underline{a}') = I - \underline{a}\ \underline{a}'$. This follows immediately from the fact that $\underline{a}\ \underline{a}'$ is idempotent because of the normalization.

2) As tr (AB) = tr (BA) and $\underline{a}'\ N'N\ \underline{a}$ is a scalar, its trace is the same scalar.

The vector $\underline{a}$ is normalized such that its Euclidean norm, $\underline{a}'\,\underline{a}$, is equal to 1, and this normalization will be used as a restriction in the Lagrange function (9) with λ as a Lagrange multiplier. This normalization is also given in (7). To obtain the optimal value of $\underline{a}$ in (9), the Lagrange function is differentiated with respect to $\underline{a}$. The optimal value of $\underline{a}$ is obtained, when:

$$\left.\begin{aligned} \frac{\partial L}{\partial \underline{a}} &= 2\,N'N\,\underline{a} - 2\,\lambda\,\underline{a} = 0 \\ \text{or} \qquad (N'N)\,\underline{a} &= \lambda\,\underline{a} \end{aligned}\right\} \qquad (10)$$

The vector $\underline{a}$ is called an eigenvector, belonging to the eigenvalue λ. When (10) is premultiplied by $\underline{a}'$, one obtains:

$$\underline{a}'\,(N'N)\,\underline{a} = \lambda\,\underline{a}'\,\underline{a} = \lambda \qquad (11)$$

The maximum of the left-hand side of (11) will be attained when the largest eigenvalue of $N'N$ is chosen. The Lagrange multiplier can be interpreted as the largest eigenvalue of $N'N$. The matrix $N'N$ has rank J and is positive definite; this means that there are J positive eigenvalues. By means of (3) and (11) it is easily seen that :

$$\underline{p}'\underline{p} = \underline{a}'\,N'N\,\underline{a} = \lambda \qquad (12)$$

Formula (4) can be interpreted as follows: $\underline{a}$ is the vector of regression coefficients of size J, in a regression analysis from $\underline{p}$ on N. This can be rewritten as:

$$N'\,\underline{p} = (N'N)^{-1}\,\underline{a} \qquad (13)$$

The vector $\underline{a}$ is called the vector of factor loadings corresponding to this eigenvalue λ, while p is the principal component. The difference between λ and the trace of $N'N$ can be computed as (see also formulae (8) and (11)):

$$\text{tr}\,(D'D) = \text{tr}\,(N'N) - \lambda \qquad (14)$$

and :

$$\text{tr } (N'N) = \lambda + \text{tr } (D'D) \tag{15}$$

λ can be interpreted as the extent to which the original data matrix N may be approximated by the principal component $\underline{p}$. Since the trace of the matrix $N'N$ is equal to the sum of its eigenvalues, it can easily be seen that the largest possible value of λ is the trace of $N'N$. This case occurs for a positive semi-definite matrix $N'N$ with only one non-zero eigenvalue. This implies that:

$$\phi = \frac{\lambda}{\text{tr } (N'N)} \tag{16}$$

falls in the range between 0 and 1.
The first principal component, defined by the largest eigenvalue of (10) and its corresponding eigenvector, is the linear function of the variables n_j $(j = 1, \ldots, J)$ with the highest variance under the given normalization condition.
Usually, the first principal component does not give a satisfactory description of the information derived from the data matrix N, especially with small values of ϕ. Next, the first principal component gets an index 1 (i.e. $\underline{p}_1$) and by means of the residuals the second and following components will be computed. The matrix of residuals, defined by (5), will be called N_2. N_2 will be used as the input-matrix for the computation of the successive components. This leads, analogous to (3), (7), (10) and (11), to the following results:

Definition : $$\underline{p}_2 = N_2 \underline{a}_2 \tag{17}$$

Normalization : $$\underline{a}_2' \underline{a}_2 = 1 \tag{18}$$

Optimality condition: $$N_2' N_2 \underline{a}_2 = \lambda_2 \underline{a}_2 \tag{19}$$

and :

$$\underline{p}_2' \underline{p}_2 = \underline{a}_2' N_2' n_2 \underline{a}_2 = \lambda_2 \underline{a}_2' \underline{a}_2 = \lambda_2 \tag{20}$$.

The relative contribution from this second component is now $\lambda_2/\text{tr } (N'N)$, and the total contribution from the first two components is $(\lambda_1 + \lambda_2)/\text{tr } (N'N)$.

By means of the newly computed matrix of residuals,

$$N_3 = N_2 - \underline{p}_2 \underline{a}_2' \tag{21}$$

and formulae (17) - (19), the subsequent principal components can be computed in the same way.

As has been stated before, the original variables are frequently correlated. This means that not each variable is equally important for explaining the variation in the data-matrix N . The consequence is that the number of relevant factors in a principal component analysis will be smaller than the J original variables. But the problem of how to select the number of independent factors has not been solved in a unique way. One possible method is to eliminate those factors, of which the corresponding eigenvalues have a rapid decline (or break). Suppose for example that the first principal component has 80 percent of the total variation of the n_{ij}'s , $\forall i,j$, and the first two components represent 95 percent of the total. In that case it would be acceptable to say that there are only two "latent" variables.
The computation of the eigenvectors is not invariant against the choice of units of measurement of the data matrix N . This can be illustrated by computing the principal components from the original data matrix N and the one where each observation deviates from its mean.
A short description of the data is given in the following section.

The COROP-data

The data concern observations on infrastructural, socio-economic and environmental endowments for 40 Dutch COROP-areas.

The socio-economic variables are:

1. fiscal income per capita
2. unemployment rate
3. wealth per capita
4. index of cost of living

The environmental variables are:

5. population density
6. size of natural environment as percentage of total regional area
7. index of industrialization related to regional area
8. index of the emission of pollutants related to regional area

The infrastructure variables are:

9. density of transportation network
10. index of cultural centres and accommodations per capita
11. index of number of schools of various types per capita
12. distance to the centre of the Netherlands
13. index of various medical services per capita.

The profile elements are represented in the following table

Table 1 Matrix with profile elements at COROP-level

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	+77.1	+26.0	+31.2	+96.6	+33.2	+6.0	+35.8	+49.3	+22.1	+31.8	+62.5	+1.1	+19.6
2	+77.9	+28.6	+33.1	+84.3	+27.1	+2.0	+18.6	+28.0	+24.1	+31.4	+90.8	+1.1	+27.3
3	+80.4	+36.5	+42.5	+88.0	+23.1	+8.9	+3.4	+35.1	+21.8	+34.5	+90.8	+1.2	+81.0
4	+77.8	+37.1	+34.1	+90.0	+31.2	+35.6	+38.9	+34.2	+21.2	+13.0	+95.8	+1.7	+32.6
5	+74.3	+33.9	+39.1	+94.7	+44.9	+10.8	+66.1	+45.7	+18.6	+100.0	+82.5	+2.3	+23.1
6	+76.1	+38.6	+38.7	+88.4	+36.5	+22.6	+44.8	+57.1	+23.4	+32.5	+82.5	+1.8	+33.2
7	+82.4	+36.6	+46.1	+90.0	+38.5	+33.0	+30.4	+39.6	+23.9	+22.0	+52.5	+1.4	+70.0
8	+74.8	+20.8	+26.9	+92.5	+34.4	+24.8	+32.8	+39.9	+25.6	+30.9	+80.0	+1.4	+26.2
9	+77.0	+43.9	+37.8	+89.2	+46.6	+43.9	+46.8	+45.3	+22.3	+31.7	+77.5	+1.8	+36.8
10	+76.8	+48.6	+35.2	+89.2	+36.5	+24.2	+38.5	+29.1	+23.7	+28.2	+90.8	+2.6	+29.3
11	+80.5	+52.4	+39.5	+87.6	+20.1	+30.4	+15.2	+15.4	+31.7	+35.8	+100.0	+2.9	+40.6
12	+77.4	+41.5	+31.7	+100.0	+16.1	+34.5	+13.3	+16.5	+33.9	+99.7	+77.5	+1.9	+31.3
13	+80.2	+75.9	+45.4	+83.5	+21.1	+100.0	+27.1	+25.3	+25.1	+12.3	+62.5	+8.0	+38.8
14	+77.6	+59.3	+44.1	+88.0	+26.8	+24.4	+26.1	+27.9	+29.9	+23.6	+65.0	+2.9	+30.6
15	+83.1	+32.1	+33.7	+83.2	+9.6	+43.2	+10.5	+9.4	+39.8	+31.4	+85.0	+7.1	+67.7
16	+76.1	+37.8	+32.1	+88.8	+24.7	+7.0	+28.6	+22.8	+34.2	+16.4	+42.5	+16.7	+23.6
17	+88.1	+83.6	+51.4	+82.5	+9.3	+37.0	+11.0	+8.3	+34.1	+24.8	+90.0	+100.0	+71.8
18	+81.1	+59.6	+31.4	+90.0	+22.0	+15.0	+42.1	+19.7	+27.5	+23.9	+52.5	+14.3	+24.0
19	+87.2	+54.4	+45.3	+76.4	+10.4	+36.1	+18.3	+12.5	+38.6	+23.3	+92.5	+100.0	+48.2
20	+87.4	+100.0	+26.1	+88.8	+5.4	+69.9	+2.0	+5.9	+46.3	+23.3	+70.0	+100.0	+37.4
21	+96.4	+98.6	+73.2	+85.8	+3.7	+94.3	+4.3	+3.6	+59.6	+48.1	+80.0	+100.0	+71.7
22	+85.5	+84.4	+21.0	+91.2	+4.3	+22.5	+3.0	+4.7	+52.2	+14.9	+55.0	+100.0	+29.5
23	+93.8	+46.1	+37.3	+93.0	+4.4	+9.4	+4.5	+3.0	+50.5	+58.3	+67.5	+100.0	+76.7
24	+95.3	+99.5	+100.0	+90.4	+4.8	+59.5	+5.9	+4.4	+54.5	+25.3	+65.0	+100.8	+66.4
25	+85.7	+51.6	+42.4	+81.5	+4.5	+29.2	+6.9	+3.7	+49.2	+18.7	+60.0	+100.0	+100.0
26	+100.0	+61.2	+62.3	+80.2	+2.0	+38.3	+2.9	+1.7	+100.0	+45.2	+82.5	+100.0	+52.1
27	+83.3	+81.2	+38.4	+79.8	+6.0	+6.4	+6.1	+4.1	+41.4	+13.0	+52.5	+100.0	+34.7
28	+83.3	+71.5	+43.4	+74.7	+11.4	+5.5	+15.9	+10.7	+28.3	+13.1	+60.0	+100.0	+32.4
29	+89.2	+52.6	+34.2	+94.7	+6.2	+19.7	+5.4	+3.3	+43.4	+22.5	+60.8	+100.0	+41.5
30	+84.0	+64.5	+35.8	+84.6	+8.8	+10.9	+6.2	+8.0	+39.8	+16.2	+57.5	+100.0	+27.5
31	+82.3	+50.6	+51.5	+95.6	+42.2	+12.8	+24.5	+15.6	+32.1	+20.2	+68.0	+3.1	+22.3
32	+84.3	+41.8	+50.7	+83.9	+27.8	+15.4	+42.5	+37.7	+36.9	+19.4	+77.5	+7.4	+27.7
33	+82.6	+34.1	+32.9	+88.2	+15.8	+29.8	+14.3	+14.1	+34.1	+20.8	+72.5	+12.5	+31.4
34	+77.2	+33.0	+33.3	+94.3	+15.3	+41.5	+16.2	+16.6	+30.5	+21.9	+70.0	+8.3	+28.3
35	+77.9	+34.5	+27.6	+78.9	+16.1	+57.7	+12.3	+16.6	+38.0	+24.5	+72.5	+4.5	+34.3
36	+79.4	+35.8	+28.6	+86.5	+13.9	+24.1	+8.4	+18.7	+36.4	+23.9	+67.5	+2.7	+34.9
37	+77.7	+38.6	+31.1	+88.5	+21.2	+46.7	+16.7	+16.9	+33.8	+19.0	+62.5	+2.4	+33.2
38	+77.0	+28.8	+29.2	+75.9	+19.9	+41.9	+19.0	+18.4	+36.8	+27.5	+65.0	+1.8	+25.8
39	+81.2	+19.9	+28.4	+87.6	+6.8	+16.9	+5.7	+5.8	+88.3	+26.3	+75.8	+1.4	+39.1
40	+84.2	+48.1	+23.6	+62.7	+100.0	+66.5	+100.0	+100.0	+13.9	+32.2	+52.5	+7.7	+14.7

1) For a precise definition as well as the sources of the data, we refer to Van Veenendaal, Regionale Welvaart in Nederland, (mimeographed), Department of Economics, Free University, Amsterdam, 1981.

Next one may determine a redefinition of the data where the observations from table 1 are normalized such that the series of observations are transformed into new data with zero mean. So all observations reflect deviations from the mean.

Tabel 2 Matrix with profile elements at COROP-level, deviating from the mean

	1	2	3	4	5	6	7	8	9	10	11	12	13
1 :	-5.2	-24.8	-7.9	+10.3	+11.9	-25.6	+13.3	+27.4	-13.8	+4.2	-8.9	-34.4	-20.6
2 :	-4.4	-21.9	-6.0	-2.0	+5.8	-29.6	-3.1	+6.1	-11.8	+3.8	+18.6	-34.4	-12.9
3 :	-1.9	-13.9	+3.4	+1.7	+1.8	-22.7	-18.3	+13.2	-14.1	+6.9	+18.6	-34.3	+40.8
4 :	-4.5	-13.3	-5.0	+3.7	+9.9	+4.0	+17.2	+12.3	-14.7	-14.6	+23.6	-33.8	-7.6
5 :	-8.0	-10.5	+0.0	+8.4	+23.5	-20.8	+44.4	+23.8	-17.3	+72.4	+11.1	-33.2	-17.1
6 :	-6.2	-11.8	-6.4	+2.1	+17.2	-9.3	+23.1	+35.2	-12.5	+4.9	+11.1	-33.7	-7.0
7 :	+0.1	-13.8	+7.0	+3.7	+17.2	+1.4	+8.7	+17.7	-12.0	-5.6	-18.9	-34.1	+29.8
8 :	-7.5	-29.6	-12.2	+6.2	+13.1	-6.8	+11.1	+18.0	-10.3	+3.3	+8.6	-34.1	-14.0
9 :	-5.3	-6.5	-1.3	+2.9	+25.3	+12.3	+25.1	+23.4	-13.6	+4.1	+6.1	-33.7	+3.4
10 :	-5.5	-1.8	-3.9	+2.9	+15.2	-5.4	+16.8	+7.2	-12.2	+0.6	+18.6	-32.9	-10.9
11 :	-1.8	+2.0	+0.2	+1.3	-1.2	-1.2	+6.5	-0.5	-4.2	+7.4	+28.6	-32.6	+1.4
12 :	-4.9	-8.9	-7.4	+13.7	-5.2	+2.9	-8.4	-5.4	-2.0	+2.1	+6.1	-33.6	-8.9
13 :	-2.1	+25.5	+0.3	-2.8	-3.2	+68.4	+5.4	+3.4	-10.8	-15.3	-8.9	-27.5	-1.4
14 :	-4.7	+8.9	+5.0	+1.7	+5.5	-7.2	+4.4	+6.0	-6.0	-4.0	-6.4	-32.6	-9.6
15 :	+0.8	-18.3	-5.4	-3.1	-11.7	+11.6	-11.2	-12.5	+3.9	+3.8	+13.6	-28.4	+27.5
16 :	-6.2	-12.6	-6.8	+2.5	+3.4	-24.6	+6.9	+0.9	-1.7	-11.2	-28.9	-18.8	-16.6
17 :	+5.8	+33.2	+12.3	-3.8	-12.3	+5.4	-10.7	-13.6	-1.8	-2.8	+18.6	+64.5	+31.6
18 :	-1.2	+9.2	-7.7	+3.7	+0.7	-16.6	+20.4	-2.2	-8.4	-3.7	-18.9	-21.2	-16.2
19 :	+4.9	+4.0	+6.2	-9.9	-13.9	+4.5	-3.4	-9.4	+2.7	-4.3	+21.1	+54.5	+0.0
20 :	+5.1	+49.6	-13.0	+2.5	-15.9	+38.3	-19.7	-16.0	+10.4	-4.3	-1.4	+64.5	-2.8
21 :	+14.1	+44.2	+34.1	-1.3	-17.5	+67.7	-17.4	-18.3	+23.7	+20.5	+8.6	+64.5	+31.5
22 :	+3.2	+34.0	-18.1	+4.9	-17.3	-9.1	-18.7	-17.2	+16.3	-12.7	-16.4	+64.5	-10.7
23 :	+11.5	-4.3	-1.8	+6.7	-16.9	-22.2	-17.2	-15.9	+14.6	+2.7	-3.9	+64.5	+36.5
24 :	+13.0	+49.1	+50.9	+4.1	-16.3	+27.9	-15.8	-17.5	+18.6	-2.3	-6.4	+64.5	+26.2
25 :	+3.4	+1.2	+3.3	-4.8	-16.8	-2.4	-14.8	-18.2	+13.3	-8.9	-11.4	+64.5	+59.8
26 :	+17.7	+10.8	+23.2	-6.1	-19.3	+6.7	-18.8	-20.2	+64.1	+17.6	+11.1	+64.5	+11.9
27 :	+1.0	+30.8	+0.7	+6.5	-15.3	-25.2	-15.6	-17.8	+5.5	-14.6	-18.9	+64.5	-5.5
28 :	+1.0	+21.1	+4.3	-11.6	-9.9	-26.1	-5.8	-11.2	-7.6	-14.5	-11.4	+64.5	-7.8
29 :	+6.9	+2.2	+4.9	+8.4	-15.1	-11.9	-16.3	-18.6	+7.5	-5.1	-11.4	+64.5	+1.3
30 :	+1.7	+14.1	-3.3	-1.7	-12.5	-20.7	-15.5	-13.9	+3.1	-11.4	-13.9	+64.5	-12.7
31 :	-0.0	+0.2	+12.4	+9.3	+23.9	-18.8	+2.8	-6.3	-3.8	-7.4	-11.4	-32.4	-17.9
32 :	+2.0	-8.6	+11.6	-2.4	+6.5	-16.2	+20.8	+15.8	+1.0	-8.2	+6.1	-28.1	-12.5
33 :	+0.3	-16.3	-6.2	-6.1	-5.3	-1.8	-7.4	-7.8	-1.8	-6.8	+1.1	-23.0	-8.8
34 :	-5.1	-17.4	-5.8	+8.0	-6.3	+9.9	-5.5	-5.3	-5.4	-5.7	-1.4	-27.2	-11.9
35 :	-4.4	-15.9	-11.5	-7.4	-5.2	-7.5	-9.4	-5.3	+2.1	-3.1	+1.1	-31.0	-5.9
36 :	-2.9	-15.4	-10.5	+0.2	-7.4	+26.1	-13.3	-3.2	+0.5	-3.7	-3.9	-32.8	-5.3
37 :	-4.6	-19.8	-8.0	-5.8	-9.1	+15.1	-6.0	-5.0	-2.1	-8.6	-8.9	-33.1	-7.0
38 :	-5.3	-21.6	-9.9	-10.4	-1.4	+10.3	-2.7	-3.5	+0.1	-0.1	-6.4	-33.7	-14.4
39 :	-1.1	-30.5	-18.7	+1.3	-14.5	-14.7	-16.0	-16.4	+22.4	-1.3	+3.6	-34.1	-1.1
40 :	+1.9	-2.3	-15.5	-23.6	+78.7	+34.9	+78.3	+78.1	-22.0	+4.6	-18.9	-27.8	-25.5

Results of the Principal Component Analysis

To start the PCA for the profile elements from table 1, we give the correlation elements between the set of $J = 13$ variables. In that way it will be possible to get an idea about the level of correlation between the variables. High correlations indicate that the set of profile elements can be reduced without loss of much information. The Pearson correlation coefficient q between variables j and k is given by (see also Brouwer and Nijkamp, 1982).

$$q = \frac{\sum_{i=1}^{40} (n_{ij}-\bar{n}_j)(n_{ik}-\bar{n}_k)}{\sqrt{\sum_{i=1}^{40} (n_{ij}-\bar{n}_j)^2 \sum_{i=1}^{40} (n_{ik}-\bar{n}_k)^2}} \qquad \forall j, k \in \{1,\ldots,13\} \qquad (22)$$

where $\bar{n}_j$ and $\bar{n}_k$ are the mean values of variable j and k, respectively. The results are represented in table 3.

Table 3 Pearson's correlation between profile elements

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1.00												
2	0.64	1.00											
3	0.57	0.53	1.00										
4	-0.17	-0.09	0.04	1.00									
5	-0.43	-0.36	-0.22	-0.15	1.00								
6	0.27	0.41	0.29	-0.26	0.01	1.00							
7	-0.40	-0.30	-0.19	-0.15	0.92	0.02	1.00						
8	-0.45	-0.40	-0.24	-0.14	0.94	0.01	0.90	1.00					
9	0.72	0.38	0.37	-0.04	-0.66	0.16	-0.62	-0.67	1.00				
10	0.04	-0.15	0.13	0.21	0.19	-0.03	0.28	0.21	0.04	1.00			
11	-0.04	-0.16	0.12	0.11	-0.05	0.07	-0.05	0.02	-0.01	0.35	1.00		
12	0.75	0.74	0.38	-0.15	-0.59	0.08	-0.51	-0.59	0.61	-0.09	-0.16	1.00	
13	0.52	0.26	0.47	0.01	-0.44	0.19	-0.48	-0.38	0.38	0.09	0.18	0.43	1.00

Relative high positive correlations can be found in variables 1-3 (the socio-economic profile) and variables 5, 7 and 8 (the environmental profile). Variable 6 has small correlations with the environmental profile. This will indicate independency between these variables. Another conclusion from table 3 will be the negative correlations between the socio-economic and environmental profile.
From these highly correlated groups of variables can be concluded that a reduced number of components in the PCA will give a good fit of the original data. The matrix of the first five loadings with the corresponding relative contributions, given the original standardized data from table 1, are given in table 4 and table 5. The results are computed by means of the steps given in formulae (17) - (21). The factor loadings are the eigenvectors corresponding with eigenvalue λ and uniquely determined by means of the unit Euclidean normalization.

Table 4 Factor loadings for 5 principal components

variable	1	2	3	4	5
1	0.466	-0.052	-0.075	-0.075	-0.156
2	0.295	0.210	0.328	0.137	-0.313
3	0.227	0.041	0.013	0.114	0.150
4	0.484	-0.160	-0.225	-0.141	-0.335
5	0.115	-0.301	0.348	-0.154	0.108
6	0.185	-0.011	0.333	0.823	0.023
7	0.118	-0.306	0.434	-0.215	0.107
8	0.118	-0.323	0.349	-0.161	0.198
9	0.208	0.139	-0.186	0.051	-0.101
10	0.157	-0.100	-0.053	-0.144	0.371
11	0.403	-0.165	-0.287	-0.012	0.013
12	0.224	0.754	0.328	-0.342	0.132
13	0.234	0.113	-0.258	0.172	0.718

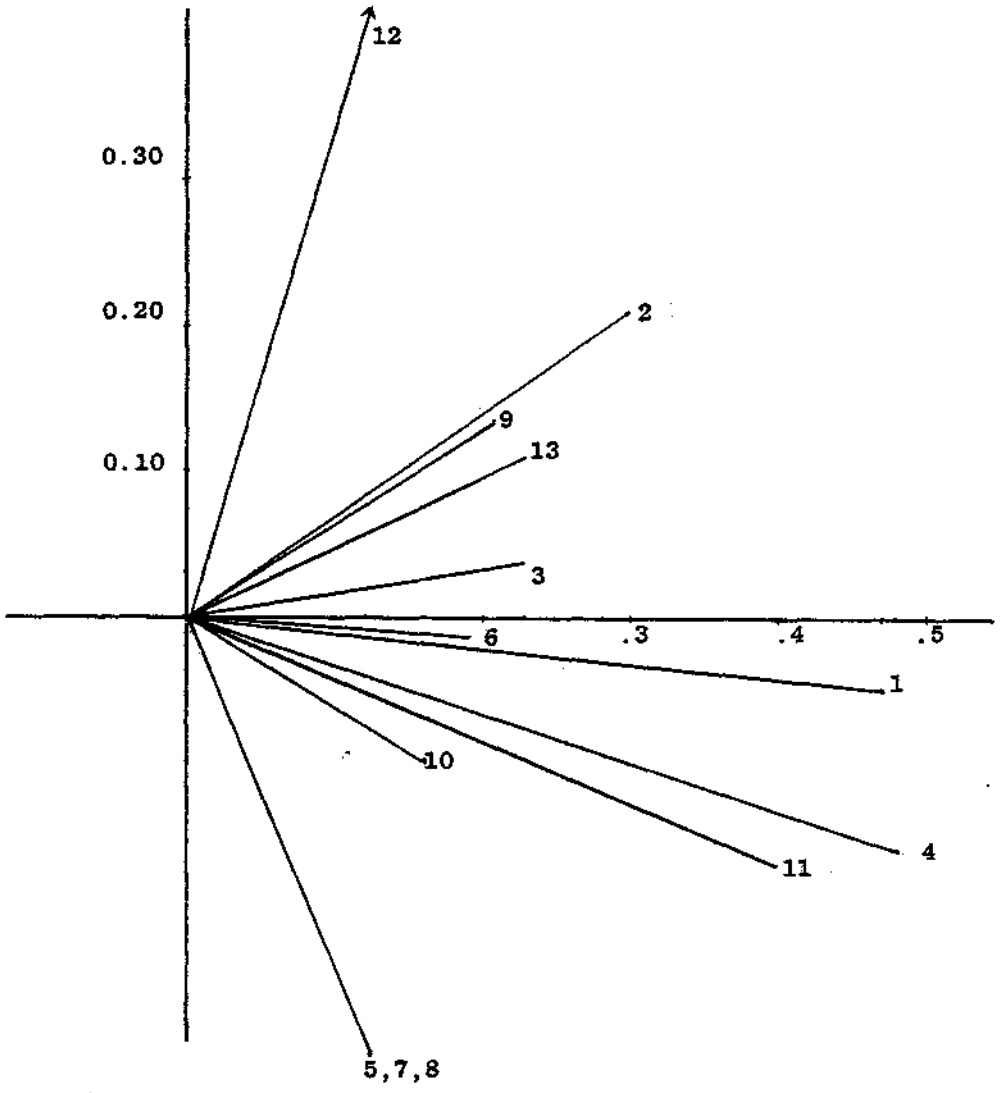
The relative contributions of the successive principal components related to these factor loadings are determined by formula (16) and are represented in table 5.

Table 5 Relative and total contribution from 5 principal components

relative contribution					total contribution	largest eigenvalue
1	2	3	4	5		
0.865	0.076	0.018	0.015	0.008	0.982	1248773

The 5 factors and their corresponding principal components explain about 98% of the information from the original data matrix. The first two factor loadings with their components have no less than 94% of the information and this two-dimensional representation is also given in figure 1.

Figure 1 Vectors of variables from the first two factor loadings



Variables 5, 7 and 8 correspond very well with each other, which has also been concluded from Pearson's correlation analysis.
In this figure, the profile elements 1 - 3 also correspond rather well.
It is easy to see that infrastructure variable 12, which represents the distance to the centre of the Netherlands has main characteristics rather independent from the other variables.

The tables 6 and 7 represented below, give the factor loadings and relative contributions, respectively, when the variables deviate from their mean. In this case, the data n_{ij}, i=1,...,40; j=1,...,13 , which are given in table 2, are not really cardinal but are measured on an interval scale.

Table 6 Factor loading when observations are deviated from the mean

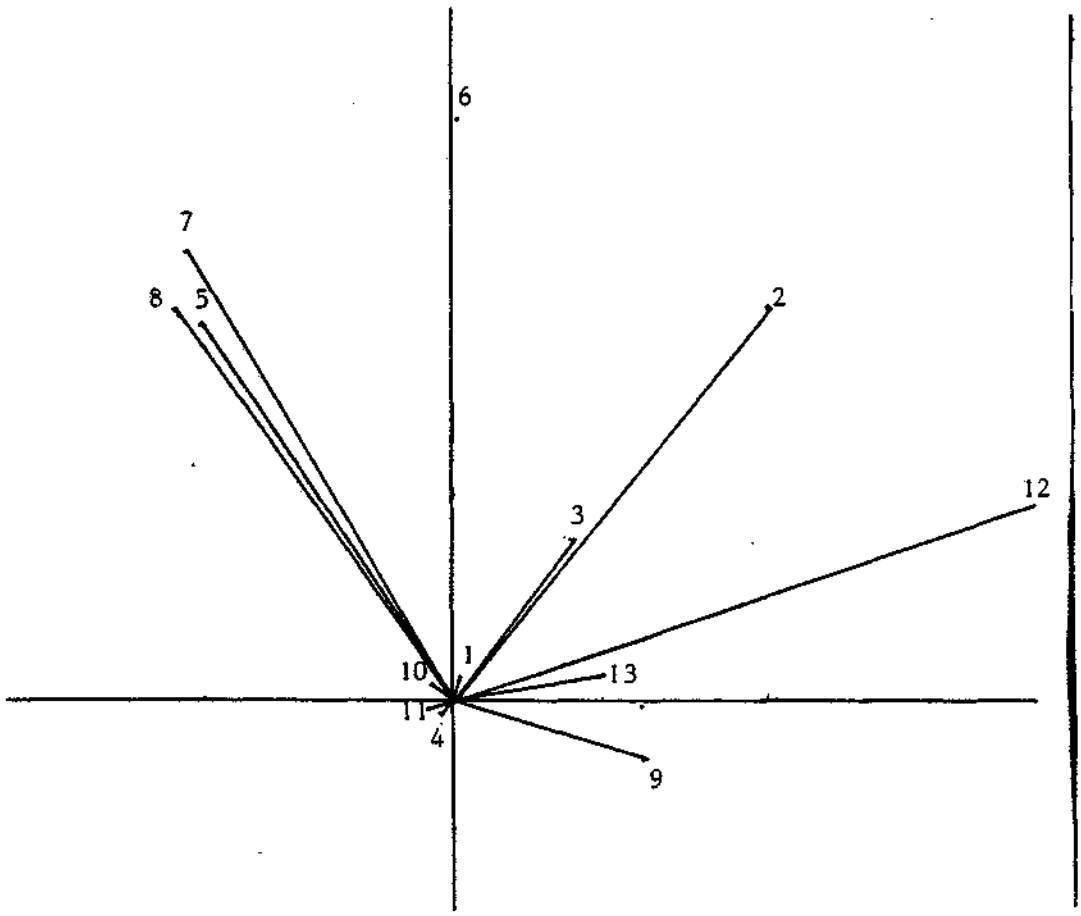
1	[illegible]	[illegible]	[illegible]	[illegible]	-0.019
2	[illegible]	[illegible]	[illegible]	-[illegible]	-[illegible]
3	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
4	-[illegible]	-[illegible]	[illegible]	[illegible]	-[illegible]
5	-0.243	[illegible]	-[illegible]	[illegible]	[illegible]
6	[illegible]	[illegible]	[illegible]	-[illegible]	-[illegible]
7	-[illegible]	[illegible]	-[illegible]	[illegible]	-[illegible]
8	-[illegible]	[illegible]	-[illegible]	[illegible]	[illegible]
9	[illegible]	-[illegible]	[illegible]	[illegible]	-[illegible]
10	-[illegible]	[illegible]	-[illegible]	[illegible]	-0.536
11	-[illegible]	-[illegible]	[illegible]	[illegible]	-0.362
12	[illegible]	[illegible]	-[illegible]	[illegible]	-[illegible]
13	[illegible]	-[illegible]	[illegible]	[illegible]	[illegible]

Table 7 Relative and total contribution with 5 principal components

Relative Contribution					Total contribution	Largest contribution
1	2	3	4	5		
0.558	0.133	0.106	0.068	0.038	0.903	120050

The 5 factor loadings and the corresponding principal components give about 90 percent of the information from the data matrix in table 2. The two-dimensional representation of the two factor loadings is given in figure 2. They have about 69% of the original data which is less than the above given 94% from table 4.

Figure 2 Two-dimensional representation of the factor loadings



This figure leads to the conclusion of two highly correlated groups of variables. On the one hand a group with variables 5, 7 and 8 and on the other hand a set with variables 1 - 3. In this case, infrastructure variable 12 will again be rather independent from the other variables.

The ordinal variant of the principal component analysis makes use of the product-moment correlation coefficient between two variables i and k. In the same way as the cardinal variant, high correlations indicate that the set of profile elements can be reduced without loss of much information. For ordinal data, Kendall's τ is a widely used correlation measure, and is defined by:

$$\tau^{ik} = \frac{S^+ - S^-}{\sqrt{(S^+ + S^- + T_i)(S^+ + S^- + T_k)}} \qquad \forall\ i,k = 1,\ldots,13 \quad (23)$$

where S^+ and S^- are the number of concordant and discordant pairs of regions, and where T_i and T_k are the number of ties (i.e., identical results) with variables i and k. The number of concordant points gives a measure of the agreement between some pair of variables, and with discordant points the reverse holds. Table 8 gives the results of Kendall's rank-correlation analysis.

Table 8 Kendall's rank-correlation analysis

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1.00												
2	0.44	1.00											
3	0.25	0.37	1.00										
4	-0.19	-0.06	-0.04	1.00									
5	-0.56	-0.30	-0.04	0.16	1.00								
6	0.12	0.11	0.02	-0.18	-0.09	1.00							
7	-0.48	-0.24	-0.01	0.06	0.73	-0.04	1.00						
8	-0.58	-0.34	-0.10	0.12	0.82	-0.07	0.70	1.00					
9	0.50	0.25	0.02	-0.14	-0.73	0.10	-0.67	-0.72	1.00				
10	-0.07	-0.16	0.03	0.10	0.11	0.08	0.03	0.13	-0.13	1.00			
11	-0.09	-0.13	0.12	-0.01	0.05	0.10	0.00	0.07	-0.08	0.37	1.00		
12	0.67	0.71	0.24	-0.22	-0.63	0.12	-0.43	-0.68	0.60	-0.28	-0.27	1.00	
13	0.40	0.21	0.28	-0.13	-0.43	0.24	-0.45	-0.40	0.33	0.07	0.19	0.24	1.00

The degree of correlation gives in this case the same conclusions as the one following from the cardinal case in table 3. A highly positive correlated group consists of variables 5, 7 and 8 and - to a slightly lesser extent - the socio-economic profile with elements 1 - 3. Variable 1 is negatively correlated with the environmental profile (i.e., variables 5, 7 and 8)

Table 9 Factor loadings with ordinal principal component analysis

Variables	1	2	3
1	0.351	0.259	0.184
2	0.261	0.427	-0.241
3	0.149	0.326	0.246
4	-0.059	-0.107	-0.151
5	-0.404	0.392	0.245
6	0.058	0.088	0.241
7	-0.361	0.377	-0.120
8	-0.405	0.300	0.251
9	0.367	-0.304	-0.142
10	-0.046	-0.106	0.308
11	0.005	-0.110	0.365
12	0.327	0.350	-0.397
13	0.282	0.006	0.472

Table 10 Relative contributions from 3 factor loadings

Relative contributions			Total contribution	Largest eigenvalue
1	2	3		
0.348	0.215	0.175	0.738	4.53

The conclusion can be drawn that the sign and pattern from the two principal components from table 9 correspond rather well with the one from table 6. The first principal component and the corresponding factor loading account for about 35% of the total variation in the transformed data matrix. Next to this there is no clear correspondence between the factor-loadings for the cardinal and ordinal case with the third principal component.

3. Grouping of variables

A very elementary analysis of clustering is the so-called linkage analysis. The linkage between variables is determined by means of either their correlation elements or the Euclidean distance measure. The sets of variables are formed by the following stepwise procedures (see also Hauer and Van der Knaap, 1973).

1. for every column in the correlation or distance matrix the maximum, resp. minimum value are determined. This means that for every variable the one with highest correlation or smallest distance is determined.
2. the first group consists of the two variables which correspond best to each other. . This will be based on the correlation or distance criterium.
3. the "nearest neighbour" is determined by the variable which correspond best with one of the above given variables. When there are no higher order nearest neighbours the same procedure starts again with a following group of variables until all variables are classified.

The correlations from table 3 are in descending order: 8-5, 5-8, 7-5, 12-1, 1-12, 9-1, 2-12, 3-1, 13-1, 6-2, 10-11, 11-10, 4-6. These sets of correlation elements immediately lead to the following group of variables.

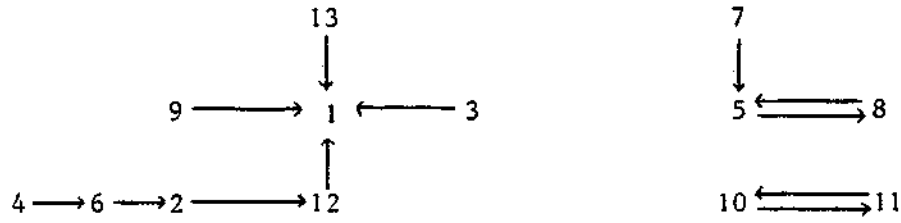


Figure 3 Classification of variables by correlation criterion.

This analysis gives three groups of variables, represented in figure 3. The groups consists of the elements {5,7,8}, {1,2,3,4,6,9,12,13} and {10,11}. The first set of elements can be interpreted as an environmental profile, and the second as a mixed type of socio-economic and infra-structural variables.

In the same way the classification may be made by means of a distance criterion for the variables. The Euclidean distance between variables j and j' is determined by:

$$d^{jj'} = \sqrt{\sum_{i=1}^{40} (n_{ij}-n_{ij'})^2} \qquad \forall\ j,j' = \{1,\ldots,13\} \qquad (24)$$

The results are represented in the following table.

Table 11 Euclidean distance elements between the variables.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	+0.00	+232.79	+283.48	+68.50	+408.66	+349.72	+409.82	+408.73	+301.39	+360.88	+119.15	+390.36	+286.92
2	+232.79	+0.00	+136.69	+271.04	+275.88	+193.29	+277.27	+281.10	+162.02	+227.24	+218.58	+225.04	+170.60
3	+283.48	+136.69	+0.00	+315.06	+195.00	+154.60	+199.67	+199.95	+106.28	+140.98	+236.11	+263.64	+114.03
4	+68.50	+271.04	+315.06	+0.00	+430.46	+382.61	+431.48	+429.77	+336.47	+383.29	+134.25	+435.47	+319.62
5	+408.66	+275.88	+195.00	+430.46	+0.00	+194.30	+47.90	+42.11	+210.93	+137.08	+348.75	+372.91	+233.03
6	+349.72	+193.29	+154.60	+382.61	+194.30	+0.00	+199.98	+198.99	+162.89	+176.87	+306.59	+309.17	+180.09
7	+409.82	+277.27	+199.67	+431.48	+47.90	+199.98	+0.00	+54.86	+217.31	+137.93	+350.45	+373.30	+242.86
8	+408.73	+281.10	+199.95	+429.77	+42.11	+198.99	+54.86	+0.00	+217.35	+141.10	+346.77	+378.72	+234.48
9	+301.39	+162.02	+106.28	+336.47	+210.93	+162.89	+217.31	+217.35	+0.00	+140.35	+259.69	+238.25	+126.73
10	+360.88	+227.24	+140.98	+383.29	+137.08	+176.87	+137.93	+141.10	+140.35	+0.00	+295.43	+310.20	+167.61
11	+119.15	+218.58	+236.11	+134.25	+348.75	+306.59	+350.45	+346.77	+259.69	+295.43	+0.00	+384.54	+240.56
12	+390.36	+225.04	+263.64	+435.47	+372.91	+309.17	+373.30	+378.72	+238.25	+310.20	+384.54	+0.00	+258.54
13	+286.92	+170.60	+114.03	+319.62	+233.03	+180.09	+242.86	+234.48	+126.73	+167.61	+240.56	+258.54	+0.00

The pairs of variables with corresponding distances in ascending order are :

8-5, 5-8, 7-5, 4-1, 1-4, 3-9, 9-3, 13-3, 11-1, 2-3, 10-5, 6-3, 12-2 .

This gives the following sets of related variables.

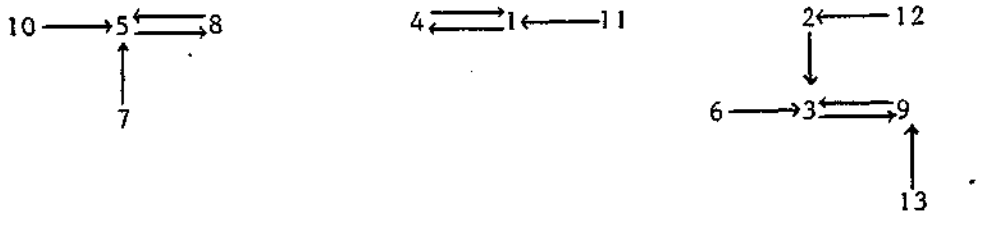


Figure 4. Classification of variables by a distance criterion.

4. Cluster Analysis

The aim of cluster analysis is to separate data into consistent groups by means of a classification procedure. It has been developed to group either individuals or variables into n subsets C_i, $i=1,\ldots,n$. When J elements are classified, the corresponding clusters have the following characteristics:

$$C_i \cap C_j = \phi \quad , \quad i \neq j$$

$$\bigcup_{i=1}^{n} C_i = \{1,2,\ldots,J\} \tag{25}$$

The intersection between two different clusters is empty and the union of the clusters consists of all grouped variables or observations. The set of clusters is determined by means of a certain criterion for partitioning the original elements. It is expected to have a high degree of diversity (dissimilarity) between each cluster and a high degree of homogeneity or agreement (similarity) within each cluster.

So the set of clusters is determined by a chosen (dis)similarity criterion. Among the most frequently used criteria are :

1. minimize the correlation (for example the product moment correlation coefficient) between the elements from different subsets.
2. maximize the correlation between the elements within a subset.
3. minimize the (Euclidean) distance between the elements within a subset.
4. maximize the (Euclidean) distance between the elements from different subsets.

The partitioning following from these criteria leads to a subset of elements which will be called clusters (see also Everitt, 1974; Hartigan, 1975). Next to the above given criteria, clustering methods can be distinguished by means of the way the elements are combined, i.e. in a hierarchical versus a non-hierarchical way.

The results of a hierarchical cluster analysis can be represented by means of a dendogram. This graphical representation of the measure of association between the elements which have been clustered, is determined by the classification criterion.
In the following figures the results of the cluster analysis with the Dutch profile elements are represented in a dendogram with the four above given classification criteria.

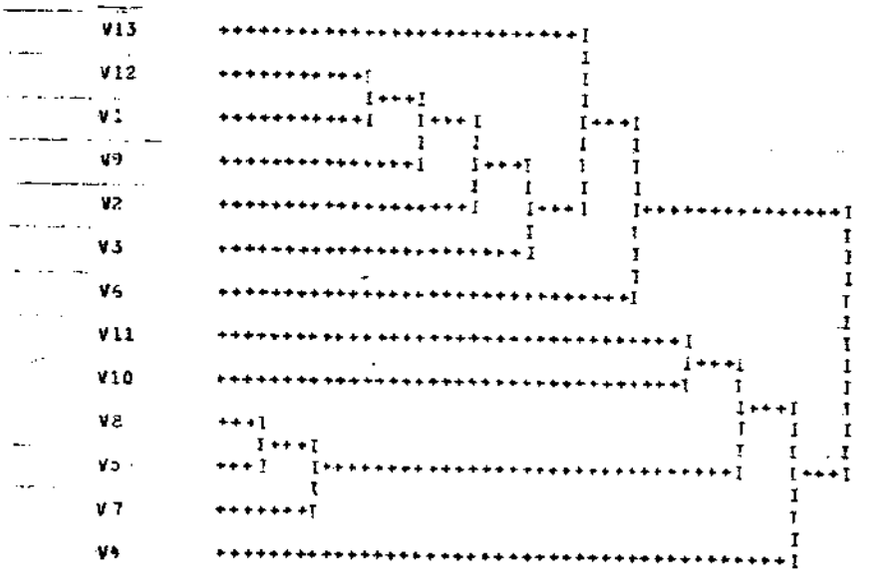


Fig.5a. Dendogram with minimization of Pearson's correlation between clusters.

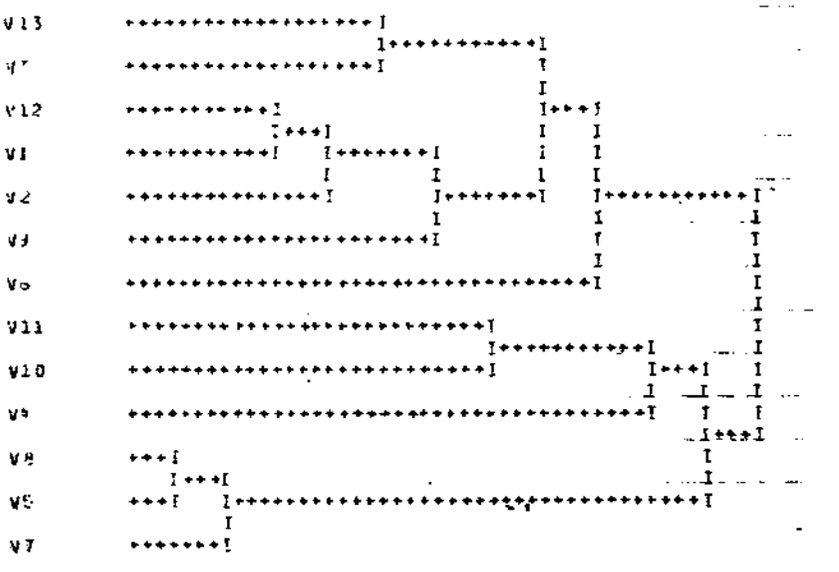


Fig.5b. Dendogram with maximization of Pearson's correlation within clusters.

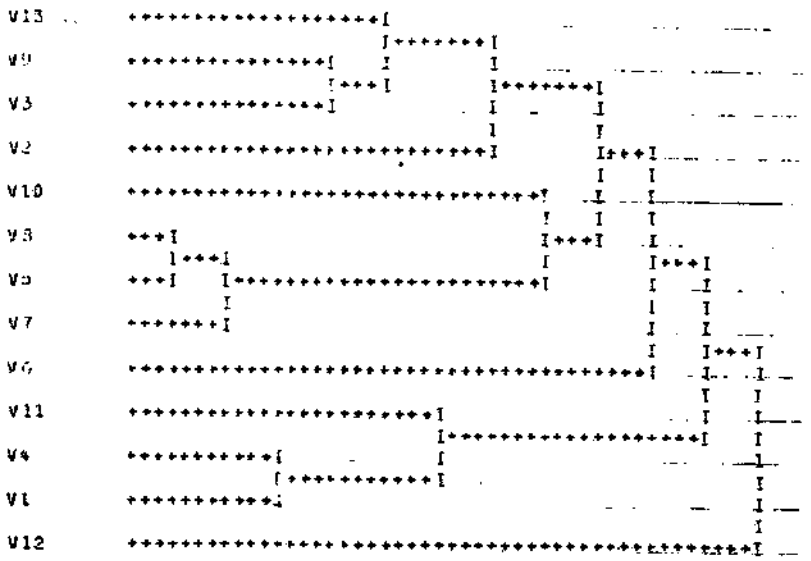


Fig.5c. Dendogram with minimization of the Euclidean distance within clusters.

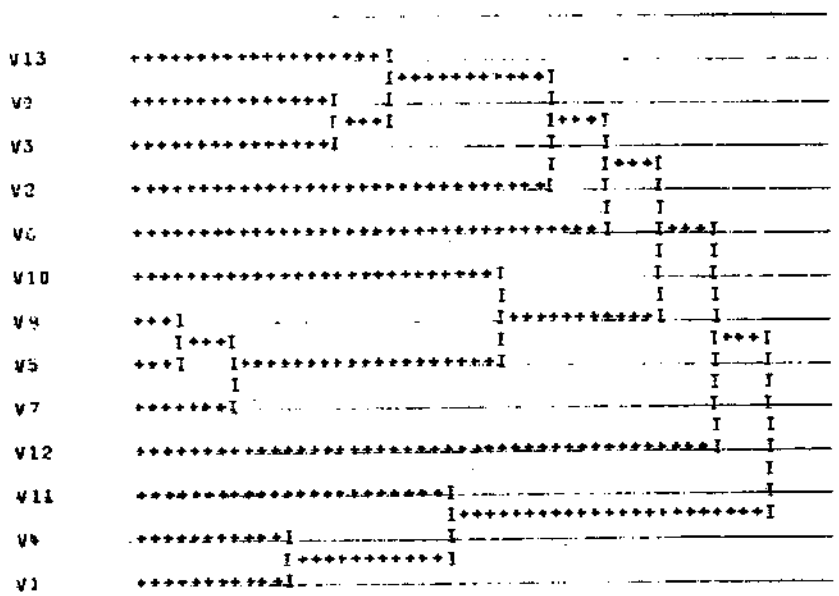


Fig.5d. Dendogram with maximization of the Euclidean distance between clusters.

When the profile elements are classified into four groups we have the following clusters with the different classification criteria:

a. $C_1 = \{5,7,8\}$

$C_2 = \{1,2,3,6,9,12,13\}$

$C_3 = \{10,11\}$

$C_4 = \{4\}$

b. $C_1 = \{5,7,8\}$

$C_2 = \{10,11\}$

$C_3 = \{1,2,3,6,9,12,13\}$

$C_4 = \{4\}$

c. $C_1 = \{1,4,11\}$

$C_2 = \{2,3,5,7,8,9,10,13\}$

$C_3 = \{6\}$

$C_4 = \{12\}$

d. $C_1 = \{1,4,11\}$

$C_2 = \{5,7,8,10\}$

$C_3 = \{2,3,6,9,13\}$

$C_4 = \{12\}$

The results from the two distance criteria correspond rather well with each other and the same holds for the two classification criteria for correlation analyses.
When the clusters are determined by a correlation criterion, the environmental profile elements are grouped rather strongly because of their high correlation values.
Because of the small correlation values of variables 10 and 11 with the other one and their rather high inter-correlation values, these two variables become a group of their own.

These results correspond rather well with the one from the previous section where the variables are grouped by means of a rather simple criterion. The same holds for the distance classification criterion. When the.

distance classification will be used in the partitioning procedure a conclusion which follows is the extreme characteristics from variable 12. Besides, the environmental profile elements correspond more or less with the socio-economic and infrastructural variables, 2, 3, 9, 10 and 13. This type of cluster analysis can be generalized to data which are measured at a nominal or ordinal scale. A useful classification-criterion may be Kendall's rank correlation elements. Next to it, a qualitative cluster analysis can be used also by means of either a multidimensional scaling method or a principal component analysis. A multidimensional scaling procedure will be developed in the following section.

5. Multidimensional Scaling Methods

There are different reduction methods for qualitative data. In previous sections, a principal component analysis and a cluster analysis with either quantitative or qualitative data has been developed. Both methods aim to reduce the original data matrix (with observations measured on a metric or non-metric scale) by means of some previously-chosen criterion. Multidimensional scaling (MDS) methods have been developed for the same reason. The final solution of an MDS procedure, i.e. the transformed data measured on a cardinal metric with a reduced number of dimensions, is not necessarily multidimensional. It depends on a 'goodness-of-fit statistic' whether the final solution is uni- or multi-dimensional. A technical and empirical survey of MDS-procedures can also be found in Kruskal (1964a, 1964b), Nijkamp (1979), Van Setten and Voogd (1978), Voogd (1978). Suppose again a data matrix N, with elements n_{ij} measured on an ordinal metric, $i=1, \ldots, I$; $j=1, \ldots, J$, where n_{ij} can be interpreted as the i-th observation with variable j . By means of scaling methods, this matrix N will be reduced to a matrix V of order $I \times K$ ($K < J$), where v_{ik} is the ith observation of attribute k , which is measured on a cardinal scale. Matrix V will be computed in such a way that the elements v_{ik} are consistent with the original elements n_{ij} from the ordinal data-matrix N. Consistency will be determined by a 'goodness-of-fit' statistic, which has to be maximized. The reduction of dimensions means that scaling methods make use of $(J - K)$ degrees of freedom, necessary to have a consistent transformation from the ordinal to cardinal information.

Suppose a symmetrix paired comparison matrix Δ, with elements $\delta_{ii'}$, ($i, i' = 1, \ldots, I, i \neq i'$), representing rank numbers, i.e.

$$\Delta = \begin{pmatrix} - & & & \\ \delta_{21} & - & & \\ \delta_{31} & \delta_{32} & - & \\ \vdots & & & \\ \delta_{I1} & \delta_{I2} & \cdots & \delta_{I,I-1} \quad - \end{pmatrix}$$

The matrix Δ contains $\frac{1}{2}I(I-1)$ relations.

Kruskal (1964a, 1964b) has hypothesized that there is a monotone relationship between the elements δ_{ij} and a distance-metric d_{ij}. By means of this relationship a 'goodness-of-fit statistic' may lead to a formal definition of the 'best' solution, in such a way that the cardinal data corresponds to the original rank numbers in an optimal way.
Kruskal's method is as follows:

$$\min_{N} \phi = f(D - D^0) \tag{26}$$

subject to :

$$D^0 \overset{M}{=} \Delta$$

$$D = g(N)$$

with :

D : a symmetric I x I matrix with elements $d_{ii'}$, representing 'distances'
N : I x J matrix
D^0 : a symmetric I x I matrix with elements d^0_{ij}
$\overset{M}{=}$: monotone relationship, i.e. :
$\delta_{ij} > \delta_{kl} \Rightarrow d^0_{ij} \geq d^0_{kl}$.

The functional form $g(N)$ is a Minkowski distance-function. A well-known metric is one which is homogeneous of order one and invariant against translations, i.e.

$$d_{ij} = \sum_{p} | n_{ip} - n_{jp} | \tag{27}$$

The Minkowski metrix can be written in a general form as :

$$d_{ij}(c) = \left\{ \sum_{p} |n_{ip} - n_{jp}|^{c} \right\}^{1/c}, \text{ with } c \geq 1 \qquad (28)$$

Formula (27(follows from (28) when c = 1. Also, the Euclidean distance function (with c = 2) is frequently used. When $c > 1$, larger differences in coordinates are assumed to have greater weights than smaller ones. With ordinal scaling procedures rank numbers will be represented in a geometric space. But it has to be noted that the solution which has been found, is an approximation to the original observations. For that reason, it will be necessary to choose an 'index-of-fit' as a selection criterion for the final solution. Kruskal (1964a) developed a 'goodness-of-fit criterion' which is scale-invariant and equal to :

$$\phi = \left(\frac{\sum_{i} \sum_{j} (d_{ij} - k(\delta_{ij}))^2}{\sum_{i} \sum_{j} d_{ij}^2} \right)^{\frac{1}{2}}, \; i > j \qquad (29)$$

ϕ will be minimized until a convergent solution is attained with a maximum value of the goodness-of-fit statistic. The abovementioned steps are summarized in the following figure which represents the algorithmic structure of an ordinal MDS-procedure.

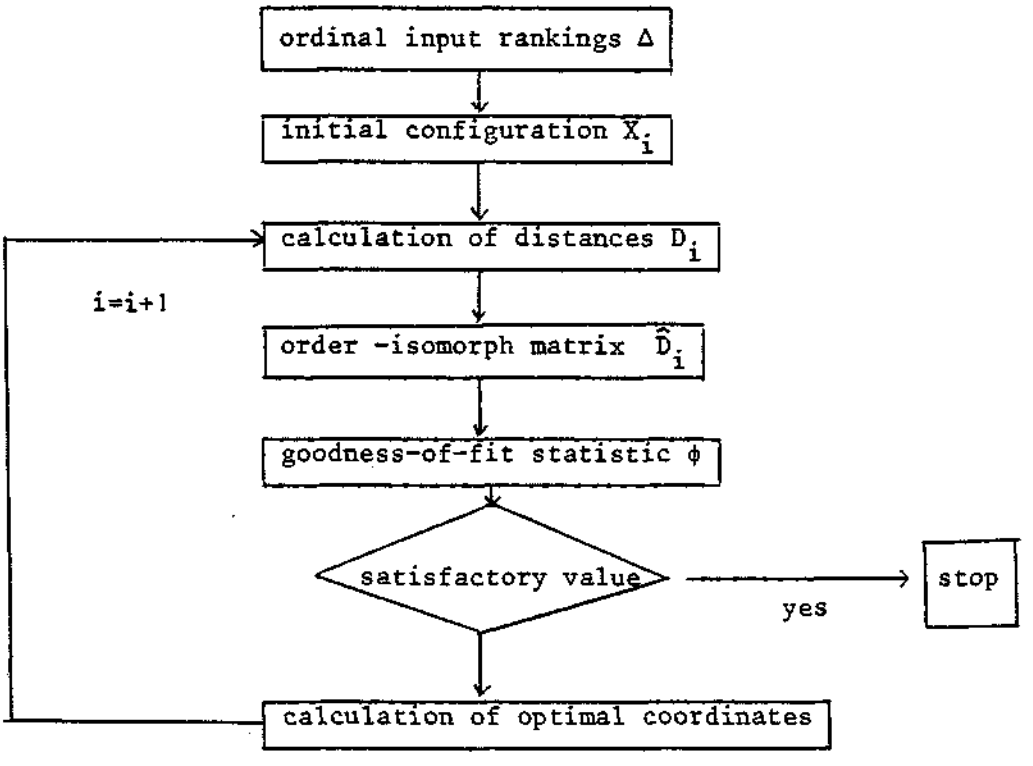


Figure 6. Algorithmic structure of an ordinal multidimensional scaling procedure.

When the observations from Table 1 are ordered from low to high for all 13 profile elements, the results of an ordinal MDS procedure can be illustrated. Table 12a contains the two-dimensional solution when an MDS-procedure is carried out with the COROP-areas and the results with the profile elements are represented in Table 12b.

1	-.2462	.1334
2	-.2519	.1198
3	-.5255	.3745
4	-.2519	.1199
5	-.5252	.3774
6	-.2517	.1203
7	-.5247	.3816
8	-.2429	.1438
9	-.2513	.1217
10	-.2520	.1199
11	-.2662	.1125
12	-.2520	.1198
13	-.5279	.3749
14	-.2520	.1199
15	-.2638	.1142
16	-.2516	.1195
17	-.1941	-.0470
18	-.5246	.3812
19	-.1614	-.0376
20	-.1116	-.0316
21	-.1352	-.0379
22	-.1055	-.0265
23	-.2808	-.0720
24	-.1034	-.0290
25	-.2793	-.0709
26	-.2740	-.0702
27	-.1096	-.0267
28	-.1052	-.0263
29	-.1320	-.0030
30	-.1085	-.0269
31	-.2520	.1200
32	-.2521	.1200
33	-.2520	.1199
34	-.2521	.1200
35	-.2520	.1199
36	-.2521	.1200
37	-.2521	.1200
38	-.2521	.1200
39	-.2520	.1199
40	.1623	1.0000

Table 12a. Two-dimensional metric representation of COROP-areas.

1	.1397	-.0207
2	.1193	.0749
3	.1192	.0748
4	.1401	-.0157
5	-.2022	-.2507
6	.1174	.0614
7	-.2024	-.2506
8	-.2022	-.2507
9	.1192	.0746
10	.1188	.0712
11	.1409	-.0141
12	-.5256	.3731
13	.1192	.0736

Table 12b. Two-dimensional metric representation of ordinal profile elements.

6. Conclusion

In this paper, some useful reduction methods have been developed which deal with either qualitative or quantitative data.

On the one hand, a principal component analysis is a reduction method which is determined by either the correlation or the covariance elements between the set of profile elements. The sign and size of the first two components of the ordinal variant and cardinal variant with observations derived from the means correspond rather well with each other (see also Tables 6 and 9).

Section 3 gives a very elementary analysis of clustering, the so-called linkage analysis. This can be seen as a first step to Section 4 where cluster analysis has been generalized and is dependent on a chosen (dis)similarity criterion. An ordinal equivalent of cluster analysis may be determined by, for example, Kendall's rank correlation analysis.

Finally, a multidimensional scaling procedure with ordinal data has been developed. This scaling procedure will lead to a reduced number of dimensions which are measured on a cardinal metric.

The numerical applications carried out with these methods show that reduction techniques may be useful tools for dealing with information measured in either a quantitative or a qualitative metric.

References

Adelman, I., and C.T. Morris, The Derivation of Cardinal Scales from Ordinal Data, Economic Development and Planning (W.Sellekaerts, ed.), MacMillan, London, 1974, pp. 1-39.

Brouwer, F., and P. Nijkamp, Multiple Rank Correlation Analysis, Research Memorandum 1982-9, Free University, Amsterdam, 1982.

Coombs, C.H., A Theory of Data, Wiley, New York, 1964.

Everitt, B.S., Cluster Analysis, Heinemann Educational Books Ltd., London, 1974.

Golledge, R.G., L.J. Hubert and G.D. Richardson, The Comparison of Related Data Sets: Examples from Multidimensional Scaling and Cluster Analysis, paper for the Regional Science Association, 1981 (mimeographed).

Hauer, J., and G.A. van der Knaap, Sociale Geografie en Ruimtelijk Onderzoek, Universitaire Pers Rotterdam, Rotterdam, 1973 (in Dutch)

Hartigan, J.A., Clustering Algorithms, Wiley, New York, 1975.

Kruskal, J.B., Multidimensional Scaling by Optimizing Goodness of Fit to a non-metric Hypothesis, Psychometrika, vol. 29, no.1, 1964a, pp. 1-27.

Kruskal, J.B., Nonmetric Multidimensional Scaling: a Numerical Method, Psychometrika, vol. 29, no.2, 1964b, pp. 115-129.

Maddala, G.S., Econometrics, McGraw-Hill, Tokyo, 1977.

Nijkamp, P., Multidimensional Spatial Data and Decision Analysis, Wiley, New York, 1979.

Nijkamp, P., and J.H. Paelinck, Operational Theory and Method in Regional Economics, Saxon House, Farnborough, 1976.

Nijkamp, P., and P. Rietveld, Soft Econometrics as a Tool for Regional Discrepancy Analysis, Research Memorandum 1981-19, Free University, Amsterdam, 1981.

Okabe, A., Statistical Analysis of the Pattern Similarity between two Sets of Regional Clusters, Environment and Planning A, vol. 13, 1981, pp. 547-562.

Setten, J. van, and J.H. Voogd, A Comparison of Optimization Methods for Ordinal Geometric Scaling Procedures, Research Centre for Physical Planning, PSC-TNO, Delft, 1978.

Theil, H., Principles of Econometrics, North-Holland Publishing Company, Amsterdam, 1979.

Veenendaal, W. van, Regional Welvaart in Nederland (Regional Welfare in the Netherlands), Free University, Department of Economics, Amsterdam, 1981, (mimeographed in Dutch).

Voogd, J.H., On the Principles of Ordinal Geometric Scaling, Research Centre for Physical Planning, PSC-TNO, Delft, 1978.

Serie Research Memoranda:

1979-1	A.A. Schreuder-Sunderman, Frans Blommaert en Hein Schreuder,	Werknemers en Soc. Jaarverslag.
1979-2	H. Schreuder,	De maatschappelijke verantwoordelijkheid van Ondernemingen.
1979-3	P. Nijkamp en P. Rietveld,	Multilevel Multi-objective Models in a Multiregional System.
1979-4	J. Arntzen, G. Bornmalm-Jardelöw and P. Nijkamp,	Duality, Segmentation and Dynamics on a Regional Labour Market.
1979-5	P. Nijkamp, A. Soffer	Soft Multicriteria Decision Models for Urban Renewal Plans.
1979-6	drs. A.J. Mathot	A Model of choosing a car with or without a credit.
1979-7	H. Blommestein, P. Nijkamp en W. van Veenendaal,	Shopping Perceptions and preferences: A multi-dimensional Attractivenerss Analysis of Consumer and Entrepreneurial Attitudes.
1979-8	H.J.Blommestein and F.C. Palm	The Aggregate Demand for Money in the Netherlands- a new look at a study of the Bank of the Netherlands.
1980-1 Jan.	P. Nijkamp and H. Voogd	New Multicriteria Methods for Physical Planning by Means of Multidimensional Scaling Techniques.
1980-2	Hidde P. Smit	Medium- and Long-Term Models for the Escap-Region - A review of existing models and a proposal for a new model system -
1980-3 april	P.v. Dijck en H. Verbruggen	Productive Employment in Developing Countries' Exporting Industries
1980-4	P. Nijkamp en L. Hordijk	Integrated Approaches to Regional Development Models; A survey of some Western European Models
1980-5	P. Nijkamp	Soft Econometric Models. An analysis of Regional Income Determinants
1980-6	P. Nijkamp en F. van Dijk	Analysis of Conflicts in Dynamical Environmental Systems via Catastrophe Theory.
1980-7 juni	E. Vogelvang	A short term econometric model for the consumer demand of roasted coffée in the Netherlands
1980-8	N. van Hulst	De effectiviteit van de geleide loonpolitiek in Nederland.
1980-9 Oct. 1980	P. Nijkamp	A survey of Dutch integrated Energy-Environmental-Economic Policy Models
1980-10 Oct. 1980	P. Nijkamp	Perspectives for Urban analyses and policies
1980-11 Oct. 1980	P. Nijkamp	New developments in multidimensional geographical data and policy analysis
1980-12	F.C. Palm, E. Vogelvang en D.A. Kodde	Efficient Estimation of the Geometric Distributed Lag Model; Some Monte Carlo Results on Small Sample Properties